PHYSICAL REVIEW B **90**, 125406 (2014)

Statistics of sub-Poissonian nucleation in a nanophase

Frank Glas*

CNRS, Laboratoire de Photonique et de Nanostructures, Route de Nozay, 91460 Marcoussis, France

(Received 16 May 2014; revised manuscript received 22 August 2014; published 5 September 2014)

We develop a fully analytical calculation of the sub-Poissonian statistics resulting from the temporal anticorrelation of the nucleation events in a supersaturated nanophase, such as occurs in particular during vapor-liquid-solid growth of nanowires [F. Glas, J. C. Harmand, and G. Patriarche, Phys. Rev. Lett. **104**, 135501 (2010)]. The sequence of nucleation events is modeled as a stochastic Markov process. The deviation from Poisson statistics is quantified by a single parameter γ, namely the ratio of the nucleation probabilities immediately after and before nucleation. We first determine self-consistently, by using q-calculus, the densities of probability of the nucleation probability, both when nucleation occurs and at an arbitrary instant. We then derive the probability for having a given number of nucleations in any given time interval. The distribution of these probabilities shows a marked narrowing with respect to Poisson statistics, in agreement with our previous experiments. We calculate explicitly the standard deviation of this distribution, which quickly saturates as the length of the time interval increases. Finally, we compute the distribution of the waiting times between nucleations. We discuss how the computed quantities vary with parameter γ. The results are in complete agreement with our numerical simulations. Somewhat surprisingly, a marked narrowing of the distribution of the numbers of nucleations occurring in fixed time intervals appears as fully compatible with a very broad distribution of waiting times.

DOI: 10.1103/PhysRevB.90.125406 PACS number(s): 81.10.Aj, 82.60.Nh, 02.50.−r, 81.07.Gf

I. INTRODUCTION

The nucleation of a new stable phase from a supersaturated parent phase is a ubiquitous phenomenon which has been studied for decades. With the rise of nanostructures, it has however recently become clear that, when the parent phase is of nanometric size, nucleation may present specific features that are not accounted for by the standard classical or atomistic theories of nucleation, which deal with macroscopic supersaturated media. Such nanosized parent phases occur in many processes, such as the catalyzed growth of semiconductor nanowires (NWs) from a liquid or solid droplet [1,2], the solid state point contact reactions between a semiconductor NW and a metal nanodot or NW [3,4], the catalyzed growth of carbon nanotubes [5], droplet epitaxy [6], and the condensation of adatoms on a nanofacet [7] or in nanoreactors [8].

One peculiar feature is that nucleation, followed by the rapid formation of a small growth unit (such as an atomic layer of limited lateral extension), may significantly deplete the nanosized parent phase and thereby alter the subsequent nucleation kinetics. This phenomenon has been identified for both vapor-liquid-solid (VLS) growth [2] and point contact growth [9] of NWs. In the present work, we develop a model and fully analytical calculations of the altered nucleation statistics. The model is generic but, to be specific, we shall phrase our discussion in the terms of VLS NW growth.

Recall that in this growth mode, a liquid (L) nanodroplet sits at the apex of the solid (S) NW and the system (often comprising a substrate on which the NWs grow epitaxially) is exposed to vapor (V) fluxes carrying the NW constituents. It is widely admitted that growth proceeds via the repeated formation of a critical two-dimensional (2D) nucleus on the top facet of the NW (in contact with the liquid), followed by the rapid completion of a solid monolayer (ML) by step flow [2,10–15]. The time needed to complete the ML after each nucleation event is indeed usually very short compared to the average time between successive nucleations [2,16,17]. Partly because this sequence of events is very fast, it has, to our knowledge, never been observed directly, but there is strong experimental evidence for it. In particular, transmission electron microscopy (TEM) experiments carried out *in situ*, during growth, show that the top facet remains at a fixed position for long times and then suddenly advances by one ML [13–15].

The LS interface has long been thought to coincide with a single crystalline plane. *In situ* TEM has however revealed that it may be constituted of a main top facet normal to the growth axis bordered by narrow truncation facets [13–15]. However, when the step bordering the growing ML reaches the edge of the top facet (either at the VLS triple phase line [12] or at its intersection with truncation facets, if any), it cannot advance anymore. For another ML to grow, a new nucleation must occur. In narrow enough NWs, the probability of a second ML nucleating before the completion of the first one is very low, and there is exactly one nucleation event per ML (mononuclear regime).

The droplet may be constituted of a foreign metal (frequently, gold) in which the NW constituents are dissolved. Conversely, in the so-called self-catalyzed growth of III–V semiconductors, the droplet contains only the NW constituents, namely a small amount of the group V element(s) dissolved in the group III element(s). In any case, the droplet is so small that the rapid formation of a ML may deplete it significantly from the NW constituents. Of course, for growth to proceed, the atoms consumed must be replaced. This happens thanks to the vapor (V) fluxes provided, usually at a constant rate, bearing in mind that, in addition to the direct impingement of vapor on the droplet, refilling may occur via other pathways, such as surface diffusion [18,19] or reemission from the neighboring surfaces [20–22].

*frank.glas@lpn.cnrs.fr

The rapid depletion of the droplet upon nucleation and ML completion has several consequences. The *in situ* TEM experiments show that the truncation facets, when present, have minimal extension just before a nucleation event and are the widest just after [14,15]. This reflects the variation of the chemical potential of the NW species in the liquid. Before nucleation, the concentrations of these species increase steadily, and so does their chemical potential. Atoms are thus transferred to the solid NW. Transfer is not to the main facet, which requires overcoming a high nucleation barrier, but to the truncation facets, which might be rough or have a low nucleation barrier [14,15]. Conversely, nucleation and ML formation deplete the liquid and reduce the chemical potential; this induces a reverse transfer to liquid from truncation facets, which therefore extend.

Before these TEM results were published, we identified the cyclic *albeit aperiodic* depletion of the droplet, correlated with nucleation and ML growth, as responsible for some remarkable statistics of the nucleation events in gold-catalyzed In(As,P) NWs [2]. Namely, we could count after growth the numbers of nucleation events occurring in successive time intervals of fixed duration T in a given NW [23]. We showed that the statistics of these numbers are not Poissonian, as would happen if the nucleation events were independent of each other (and ML completion very fast), but markedly *sub-Poissonian*. Actually, the standard deviation $\sigma(T)$ of their distribution is much smaller than the Poissonian deviation and, instead of increasing as the square root of the mean number of events in the interval, it rapidly saturates. This shows that the nucleation events are temporally anticorrelated. We interpreted this in terms of nucleation probability (NP): after a nucleation event, the concentrations and the chemical potential suddenly decrease, and so does the NP. Hence nucleation is less likely after a nucleation event (and for some time) than just before. We termed this effect "nucleation antibunching" and the corresponding growth mode "self-regulated growth."

Thanks to our calculation of the chemical potential of (Au,III,V) liquids [24], we could simulate numerically long time sequences of random nucleation events and calculate their statistics [2], in particular the standard deviation $\sigma(T)$. Each simulation assumes values of the unknown quantities entering the model (average concentrations in the liquid, effective edge energy of the 2D nucleus [12]). Since $\sigma(T)$ is the only reliable output of our experiments, we could not determine these quantities independently, but we found perfectly reasonable combinations thereof allowing us to reproduce the observed standard deviations and their saturation as T increases [2]. An example of these fitting simulations is shown in Fig. 1. We also argued that, since the concentration of group V atoms in the droplet is much lower than that of the group III atoms (this is true for both self-catalyzed growth [22] and growth catalyzed by a foreign metal), whereas the solid formed is stoichiometric, the group V concentration is much more affected by sudden ML growth than the group III concentration, so that, in turn, its variations affect much more the NP. Indeed, there is very little difference between simulations taking into account both constituents and simulations considering only the group V atoms [2]. Finally, we pointed out that, *at given average liquid composition*, the effect is all the more marked when the NW

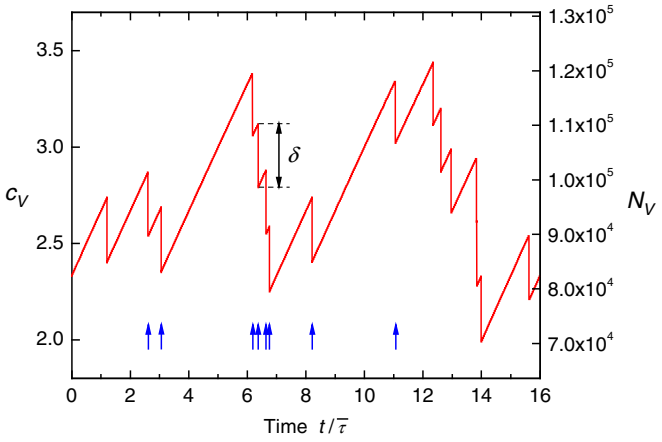

FIG. 1. (Color online) Short excerpt from a simulation that reproduces the standard deviations σ measured for In(As,P) NWs [2]: variation with time (normalized to the average time $\bar{\tau}$ between nucleation events, with arbitrary origin) of the group V concentration (left scale) and of the number of group V atoms in the droplet. This number decreases abruptly by δ after each nucleation (a sequence of which is marked by arrows) and increases linearly in between.

radius R is small, since the ML volume scales as R^2 whereas that of the droplet scales as R^3.

It immediately appears that, although simulations such as those of Fig. 1 accurately reproduce the marked narrowing of the distribution of the numbers of nucleations occurring in fixed time intervals [2], there is still a considerable randomness in the distribution of the waiting times between nucleation events. In particular, very short and very long waiting times are common. The concentrations at which nucleation occurs are also broadly distributed (in Fig. 1, they nearly vary from simple to double). One of the purposes of this work is to discuss quantitatively the interplay of self-regulation and randomness in these systems.

To this end, we develop a full *analytical* calculation of the statistics of nucleation antibunching, using only a few simplifying assumptions. Our stochastic model is described in Sec. II and the calculations are performed in Secs. III (densities of probability of the NP), IV (probability for a given number of nucleations in a time interval), and V (waiting times). Section VI is devoted to a general discussion of sub-Poissonian nucleation statistics based on our results.

II. MODEL

We consider a parent nanodroplet from which a NW of fixed radius (or, more generally, a solid of limited lateral extension) grows via 2D nucleation, at a given temperature. We are interested in cases where the amount (in the droplet) of the single NW constituent (case of a monoatomic NW) or of one of the NW constituents (polyatomic NW) is significantly affected by each nucleation event and where, in turn, its variations affect significantly the NP. As discussed above, this holds in particular for III–V compound semiconductors. Our main hypotheses are the following.

(H1) The state of the system is entirely determined by the stochastic variable $N(t)$, the number of atoms of the relevant

species in the droplet at time t. The evolution of the system is also determined by this variable, since the NP depends only on $N(t)$. The process is thus Markovian.

(H2) Nucleation and ML formation *instantaneously* decrease number N by a fixed amount $\delta > 0$, whereas refilling proceeds at a constant rate b (Fig. 1). As a consequence, the *average* waiting time between nucleations is $\bar{\tau} = \delta/b$.

(H3) For the constituent considered to have an impact on nucleation, we assume, in line with our previous work [2], that the number of this type of atoms in the liquid is low enough to be significantly diminished by the formation of a ML, but nevertheless that it never falls below the number δ of atoms contained in a ML, at least whenever nucleation occurs (Fig. 1). Would it not be the case, the system would enter another regime where the completion of the ML can no longer be taken as instantaneous, but is instead limited by the arrival rate of the atoms from the vapor to the droplet. While well worth investigating, this regime is beyond the scope of the present calculations.

We now specify the dependence of the NP P on $N(t)$. At temperature θ, P is dominated by factor $\exp[-\Delta G/(k_B\theta)]$, with ΔG the nucleation barrier and k_B Boltzmann's constant [25]. ΔG is inversely proportional to the difference of chemical potential $\Delta\mu$ between liquid and solid. $\Delta\mu$, taken here per atom (or per pair in the case of a compound semiconductor) depends on temperature and droplet composition in a known fashion [22,24,26]. Our simulations for III–V NWs [2] show that we can linearize these variations around a reference (group V) composition, unless the concentration is very small (namely, the results with and without linearization are very similar). However, we retain the exponential function, since the fast variation of the nucleation probability with $\Delta\mu$ is an essential feature of nucleation-mediated growth in general [25] and of the effect studied here in particular [2]. This also insures that P remains positive. In these conditions, the variations of the NP with N follow the simple law [2]

$$P(N) = P(N_0)\exp[A(N - N_0)], \tag{1}$$

where parameter $A > 0$, which is given explicitly in Ref. [2] for VLS NW growth, gathers factors such as nucleus geometry and edge energy, temperature, mean concentration in the liquid, and mean value of $\Delta\mu$ and of its derivative with respect to concentration. Clearly, in Eq. (1), N_0 can be chosen arbitrarily in the range where the exponential variation of the NP with N is a valid approximation. It will thus come as no surprise that none of the quantities calculated in the following depend on N_0.

According to our hypotheses, the NP varies continuously and deterministically between nucleations events. These events occur randomly at instants t_n, according to the value of the NP immediately *before* nucleation, $P(t_n)$. Nucleation causes the NP to decrease instantaneously to value $P(t_n)\exp(-A\delta)$. In the following, we use the same notation for the NP considered as a known function of N and as a stochastic variable $P(t)$.

We assume that refilling of the parent nanosized reservoir proceeds at a constant rate [hypothesis (H2)]. This means that we treat a stationary growth regime and not the transients that may occur at the beginning of growth (due to a gradual increase of mother phase supersaturation [12] or possible elastic effects induced by a substrate). We also ignore the desorption of the NW species from the liquid droplet, which may be significant for group V elements [22]. Sibirev *et al.* recently discussed the effect of desorption on nucleation statistics [27]. In common with refilling and contrary to ML formation, desorption is a continuous process. Strictly speaking, it cannot be modeled simply as a uniform decrease of refilling rate b, since it tends to depend exponentially on $\Delta\mu$ and therefore on N [22,27]. However, as a first approximation, it can be taken into account by replacing b by $b - \overline{N}/t_d$, where t_d is the characteristic desorption time [27] and $\overline{N}$ some mean value of N.

Let us stress again that, although we phrase the problem in terms of nucleation and growth of NWs, our model can describe other situations. For instance, the variations of composition and nucleation barrier induced by nucleation itself in point contact reactions are very similar [9]. The key point is that nucleation significantly and suddenly diminishes the nucleation probability, whereas refilling the parent phase is a relatively slow process.

III. DENSITIES OF PROBABILITY OF NUCLEATION

A. Definitions

The basic ingredient of our calculation is the NP $P(t)$. This is actually a density of probability per unit time, such that the probability for a nucleation event occurring within time interval $[t,t+dt]$ is $P(t)dt$. However, in order to calculate the statistics of nucleation, we shall compute other, more abstract, densities of probability. These are defined as follows.

At any given instant of the process, the number of NW atoms in the droplet is well defined, depending on the past history of nucleations in the system (and of its initial state), and the NP follows directly, via Eq. (1). Now, we may record, over the whole process, the successive values of the NP and consider their distribution. Because nucleations occur randomly, the NP is also a stochastic variable. We can thus define the density of probability of the NP, $\pi^\star(P)$, such that, *at any arbitrary instant in the process*, the probability that the NP lies within interval $[P,P+dP]$ is $\pi^\star(P)dP$. By so doing, we do not distinguish those instants where nucleation actually occurs. Obviously, at these particular instants, the NP is on average larger than over the whole process. The statistics of the NP at these instants will be characterized by another density of probability of the NP, $\widetilde{\pi}(P)$, such that, *at any instant where a nucleation event takes place*, the probability that the NP lies within interval $[P,P+dP]$ is $\widetilde{\pi}(P)dP$. We shall actually compute $\widetilde{\pi}$ first (Sec. III B), and then deduce $\pi^\star$ from it (Sec. III C). Note that both densities have the dimension of a time.

B. Density of probability $\widetilde{\pi}$ of the nucleation probability when nucleation occurs

Let us first assume that a nucleation event occurs at time t_0, when the NP is P_0. After that, as long as no new nucleation occurs, the NP varies with elapsed time τ as $P(t_0+\tau) = P_0\exp[A(b\tau - \delta)]$. We introduce the density of probability $\pi(\tau|P_0)$ for the next nucleation to occur after time τ, conditional to the fact that the first one occurred when the

NP was P_0. We have

$$\pi(\tau|P_0) = P_0 e^{A(b\tau-\delta)} \exp\left\{-\int_0^\tau P_0 \exp[A(bt-\delta)]dt\right\}, \tag{2}$$

where the first two factors account for the probability of nucleation per unit time exactly after time τ and the second exponential for the probability that no nucleation occurred between t_0 and $t_0+\tau$.

There is a one to one correspondence between τ and the NP P at time $t = t_0+\tau$, namely $P = P_0 e^{A(b\tau-\delta)}$. Hence density $\pi(\tau|P_0)$ translates into a density of probability $\pi(P|P_0)$ for the next nucleation to occur when the NP has value P (within dP), conditional to the fact that the first one occurred when the NP was P_0, according to

$$\pi(\tau|P_0)d\tau = \pi(P|P_0)dP. \tag{3}$$

By calculating the integral in Eq. (2) and differentiating the relation between τ and P, we find

$$\pi(P|P_0) = \frac{1}{Ab}\exp\left(\frac{P_0 e^{-A\delta}-P}{Ab}\right). \tag{4}$$

It is easily checked that $\int_0^\infty \pi(\tau|P_0)d\tau = 1$ and $\int_{P_0e^{-A\delta}}^\infty \pi(P|P_0)dP = 1$. Here, we note that nucleation can happen when the NP is P only if previous nucleation occurred when NP was P_0 such that $P \geqslant P_0 e^{-A\delta}$ (the NP just after previous nucleation) and because, in the absence of new nucleation, the NP can only increase. Since $\widetilde{\pi}(P)$ is defined as the density of probability for nucleation to occur, *over the whole process*, when NP equals P, one has

$$\begin{aligned}\widetilde{\pi}(P) &= \int_0^{Pe^{A\delta}} \pi(P|P_0)\widetilde{\pi}(P_0)dP_0 \\ &= \frac{e^{A\delta}}{Ab} e^{\frac{-P}{Ab}} \int_0^P e^{\frac{p_0}{Ab}}\widetilde{\pi}(p_0 e^{A\delta})dp_0,\end{aligned} \tag{5}$$

where the second equation is obtained by changing variable P_0 into $p_0 = P_0 e^{-A\delta}$. Integral equation (5) may be rewritten as a differential equation:

$$\frac{d\widetilde{\pi}}{dP}(P) = -\beta\widetilde{\pi}(P) + \gamma^{-1}\beta\widetilde{\pi}(\gamma^{-1}P), \tag{6}$$

with $\beta = \frac{1}{Ab}$ and $\gamma = e^{-A\delta} < 1$. In addition, we obviously have $\widetilde{\pi}(0) = 0$. γ, which will emerge as the fundamental parameter of our model, is simply the ratio of the NPs immediately after and before nucleation which, in the model, is independent of N [Eq. (1)].

Equation (6) belongs to the family of "pantograph equations," studied in detail by Kato and McLeod [28]. Rephrasing the results in terms of our parameters, these authors show that, if $\gamma < 1$, there is no single solution of the equation, even if the value of the function at some point is fixed. Instead, different solutions exist with different asymptotic behaviors for $P \to \infty$. However, theorem 9 of Ref. [28] states the following: (i) for any arbitrary constant L, there is a solution y_L of Eq. (6) decaying like $L\,e^{-\beta P}$ for $P \to \infty$; (ii) there is no solution y of Eq. (6), apart from the constant multiples of y_L, such that $y = o(P^{-1})$.

Our solution of Eq. (6) must be $o(P^{-1})$, otherwise we could not normalize $\widetilde{\pi}$, and therefore be of the form y_L given by Kato and McLeod, namely

$$\widetilde{\pi}(P) = L\,e^{-\beta P}\left\{1+\sum_{n=1}^\infty \frac{\gamma^{-n}\exp[\beta(1-\gamma^{-n})P]}{(1-\gamma^{-1})(1-\gamma^{-2})\cdots(1-\gamma^{-n})}\right\}, \tag{7}$$

where L is a constant to be determined.

In the following, we use the language and results of q-calculus [29], where the quantity conventionally noted q is taken equal to γ. In particular, we introduce the corresponding q-Pochhammer symbol $(a;\gamma)_n = (1-a)(1-a\gamma)(1-a\gamma^2)\cdots(1-a\gamma^{n-1})$ for $n \geqslant 1$ [with $(a;\gamma)_0 = 1$] and its limit $(a;\gamma)_\infty$ for $n \to \infty$. Then, Eq. (7) is rewritten as

$$\widetilde{\pi}(P) = L\sum_{n=0}^\infty g_n \exp(-\beta\gamma^{-n}P), \tag{8}$$

where, for $n \geqslant 0$,

$$g_n = \frac{(-1)^n \gamma^{\frac{n(n-1)}{2}}}{(\gamma;\gamma)_n}. \tag{9}$$

We will make repeated use of a result of Euler [Eq. (6.189) in Ref. [29] and Eq. (2.2.6) in Ref. [30]] which, in our terms, is written as

$$\sum_{n=0}^\infty g_n z^n = (z;\gamma)_\infty. \tag{10}$$

In particular, since $(1;\gamma)_n = 0$ for any $n \geqslant 1$, we have $(1;\gamma)_\infty = 0$ and thus [see also Eq. (2.1.1) in Ref. [31]]

$$\sum_{n=0}^\infty g_n = 0. \tag{11}$$

We find L by normalizing $\widetilde{\pi}$ according to $\int_0^\infty \widetilde{\pi}(P)dP = 1$. This relation is only approximate, since the very small values of P could only be attained if the number N of atoms became negative, which is clearly unphysical. However, because of the exponential variation of the NP with N, the low values of P are very unlikely and the contribution of these values to the integral is small enough to be neglected. This is consistent with the hypothesis made above that the number of atoms in the droplet remains larger than δ. Then

$$L = \beta\left(\sum_{n=0}^\infty \gamma^n g_n\right)^{-1} = \frac{\beta}{(\gamma;\gamma)_\infty}, \tag{12}$$

where the second equality follows from Eq. (10). Finally,

$$\widetilde{\pi}(P) = \frac{\beta}{(\gamma;\gamma)_\infty}\sum_{n=0}^\infty g_n \exp(-\beta\gamma^{-n}P). \tag{13}$$

Equation (11) proves that boundary condition $\widetilde{\pi}(0) = 0$ is satisfied.

C. Density of probability $\pi^\star$ of the nucleation probability at an arbitrary instant

We shall first derive a relation between densities $\widetilde{\pi}$ and $\pi^\star$. Let t be an arbitrary instant, in the sense that, at variance with Sec. III B, we do not suppose that a nucleation event occurs at t. If the NP at time t is P, the probability that a nucleation will occur in interval $[t,t+dt]$ is simply $P\,dt$. In addition, the probability that the NP lies within interval $[P,P+dP]$ is, by definition, $\pi^\star(P)dP$, since t is totally arbitrary. Hence the total probability that a nucleation will occur within time interval dt is $\int_0^\infty \pi^\star(P)P\,dP\,dt$. Consequently, the probability that, *if nucleation occurs within* dt, the NP is within $[P,P+dP]$ is

$$d\mathcal{P} = \frac{\pi^\star(P)dP\;P\;dt}{\int_0^\infty \pi^\star(P')P'd\,P'dt}. \tag{14}$$

But, by definition, this quantity, which is independent of time (as it should be, since t is arbitrary), is precisely $\widetilde{\pi}(P)dP$. Hence

$$\pi^\star(P) = C\,P^{-1}\widetilde{\pi}(P), \tag{15}$$

with $C = \int_0^\infty \pi^\star(P')P'dP'$. Integral C may be calculated by normalizing density $\pi^\star$. From $\int_0^\infty \pi^\star(P)dP = 1$, we find

$$\begin{aligned} C^{-1} &= \int_0^\infty \frac{\widetilde{\pi}(P)}{P}dP \\ &= \frac{\beta}{(\gamma;\gamma)_\infty}\int_0^\infty \frac{dP}{P}\sum_{n=0}^{\infty} g_n \exp(-\beta\gamma^{-n}P). \end{aligned} \tag{16}$$

Using Eq. (A19) of the Appendix (Sec. 5) with $F_n = 1$ for $n \geqslant 0$, we obtain $C^{-1} = -\beta\,\ln\gamma$ and finally

$$\pi^\star(P) = -\frac{1}{(\gamma;\gamma)_\infty \ln\gamma}P^{-1}\sum_{n=0}^{\infty} g_n \exp(-\beta\gamma^{-n}P). \tag{17}$$

D. Discussion

We have now calculated the key quantities for the rest of this study, namely the probability densities $\widetilde{\pi}$ [Eq. (13)] and $\pi^\star$ [Eq. (17)]. These densities depend only on γ ($0 \leqslant \gamma \leqslant 1$) and β, which thus appear as the primary parameters of the model. In terms of these parameters, the average time between nucleations is written as $\bar{\tau} = -\beta\,\ln\gamma$. On the other hand, parameters δ (the number of atoms in a ML) and b appear only via the primary parameters, and N_0 does not appear at all. Moreover, by changing variable P to $Q = \beta P$, we may replace $\widetilde{\pi}$ and $\pi^\star$ by reduced probability densities, functions of Q that depend only on γ, not on β. γ thus appears as the key parameter of the model. Recalling that γ is the ratio of the NPs immediately after and before nucleation (equal to 1 in the Poisson case, since the NP is then constant), γ indeed quantifies the deviation from the Poissonian case: the smaller γ, the larger the deviation.

Formally, our model can be applied for any value of γ. It is however useful to discuss which kind of values γ may assume in real systems. We recalled in Sec. I that, in the case of Au-catalyzed In(As,P) NWs, we could, by simulating long growth sequences (Fig. 1), reproduce quantitatively the variations of the standard deviation σ of the number of nucleations occurring in given time intervals [2]. The combinations of growth parameters mentioned in Sec. I translate into a value of γ, and all combinations that fit the experiments correspond to a value of about 0.70. We will indeed show in Sec. IV C 2 that, in our model, σ depends only on γ. More recently, we also studied in detail the growth kinetics of self-catalyzed GaAs NWs [21,22]. In this case, we did not systematically investigate the nucleation statistics, but we could determine not only the few physical parameters that specify our model of self-catalyzed growth [22] (in particular, the edge energy of the 2D nucleus) but also, *a posteriori*, the As concentration c_{As} in the droplet for each of the particular NWs investigated in Ref. [21]. γ is easily calculated from the model parameters and c_{As}. The latter was shown to vary depending on the As vapor flux and these variations have a large impact on γ, which depends directly on c_{As} but also indirectly via the chemical potential. We find that, over the wide range of As fluxes investigated [21,22], γ varies between about 0.06 and 0.32.

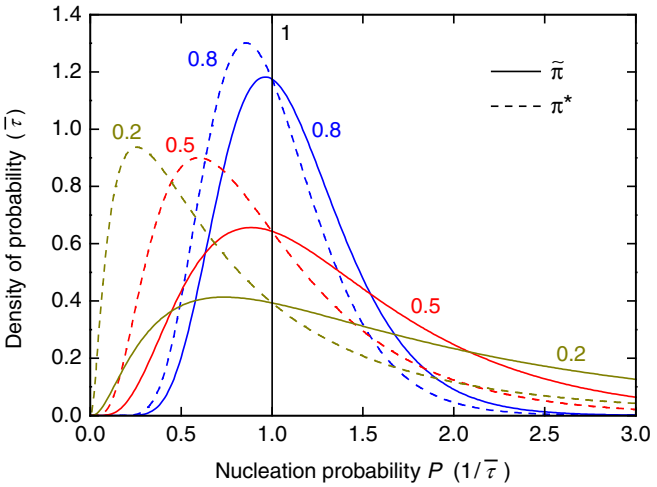

FIG. 2. (Color online) Densities of probability of the nucleation probability P (per unit of P) over the whole process ($\pi^\star$) and at those instants when nucleation occurs ($\widetilde{\pi}$), for the Poisson process ($\gamma = 1$) and for moderately ($\gamma = 0.8$) to strongly ($\gamma = 0.2$) sub-Poissonian processes. The value of γ is indicated near each curve. The time unit is $\bar{\tau}$, the average time between nucleations.

The variations of densities $\widetilde{\pi}$ and $\pi^\star$ with NP are illustrated in Fig. 2 for weakly ($\gamma = 0.8$) to strongly ($\gamma = 0.2$) sub-Poissonian processes. The Poissonian case obviously corresponds to a Dirac peak for each density, located at $P = 1$. For this figure, we have indeed chosen a time unit equal to $\bar{\tau}$, i.e., $\beta = -1/\ln\gamma$. The densities for other values of β are easily found via the change of variable mentioned above.

With respect to $\pi^\star$, $\widetilde{\pi}$ is shifted to higher P values, and the more so that the process is sub-Poissonian. This simply manifests that nucleation tends to occur at relatively high values of the NP. More precisely, Eq. (15) states that $\pi^\star/\widetilde{\pi} = C/P$. $C = \bar{\tau}^{-1}$ is the average value $\langle P\rangle$ of the NP at wholly arbitrary instants. Hence $\pi^\star > \widetilde{\pi}$ for $P < \langle P\rangle$, whereas $\widetilde{\pi} > \pi^\star$ for $P > \langle P\rangle$ (see Fig. 2, where $\langle P\rangle = 1/\bar{\tau} = 1$). In turn, using Eqs. (13) and (10), the average value of the NP at

those instants when nucleation occurs is

$$\begin{aligned}\langle P\rangle_{\rm nucl} &= \int_0^\infty \widetilde{\pi}(P) P\, dP \\ &= \frac{\beta}{(\gamma;\gamma)_\infty}\sum_{n=0}^{\infty} g_n \int_0^\infty P\, \exp(-\beta\gamma^{-n}P) dP \\ &= \frac{\beta}{(\gamma;\gamma)_\infty}\sum_{n=0}^{\infty} g_n(-\beta\gamma^{-n})^{-2} \\ &= \frac{1}{\beta(1-\gamma)} = -\frac{\ln\gamma}{1-\gamma}\frac{1}{\bar{\tau}}. \end{aligned} \tag{18}$$

As expected, the ratio $\langle P\rangle_{\rm nucl}/\langle P\rangle = -\ln\gamma/(1-\gamma)$ is always larger than 1 (since $\gamma < 1$) and tends to 1 when the process becomes Poissonian.

The corresponding standard deviations can also be calculated easily. Indeed, using Eq. (17) and then Eq. (10),

$$\begin{aligned}\int_0^\infty \pi^\star(P) P^2 dP &= -\frac{1}{(\gamma;\gamma)_\infty \ln\gamma}\sum_{n=0}^{\infty}\frac{g_n}{\beta^2\gamma^{-2n}} \\ &= -\frac{(\gamma^2;\gamma)_\infty}{\beta^2(\gamma;\gamma)_\infty\ln\gamma} = -\frac{1}{\beta^2(1-\gamma)\ln\gamma}, \end{aligned} \tag{19}$$

so that the variance of $\pi^\star$ is

$$\sigma^2_{\pi^\star} = -\frac{1}{\beta^2(1-\gamma)\ln\gamma} - \frac{1}{\beta^2(\ln\gamma)^2} = -\left(1+\frac{\ln\gamma}{1-\gamma}\right)\frac{1}{\bar{\tau}^2}. \tag{20}$$

Similarly, using Eq. (13) and then Eq. (10),

$$\begin{aligned}\int_0^\infty \widetilde{\pi}(P) P^2 dP &= \frac{\beta}{(\gamma;\gamma)_\infty}\sum_{n=0}^{\infty}\frac{2g_n}{\beta^3\gamma^{-3n}} = \frac{2}{\beta^2}\frac{(\gamma^3;\gamma)_\infty}{(\gamma;\gamma)_\infty} \\ &= \frac{2}{\beta^2(1-\gamma)(1-\gamma^2)}, \end{aligned} \tag{21}$$

so that the variance of $\widetilde{\pi}$ is

$$\begin{aligned}\sigma^2_{\widetilde{\pi}} &= \frac{2}{\beta^2(1-\gamma)(1-\gamma^2)} - \frac{1}{\beta^2(1-\gamma)^2} \\ &= \frac{1}{\beta^2(1-\gamma^2)} = \frac{(\ln\gamma)^2}{1-\gamma^2}\frac{1}{\bar{\tau}^2}. \end{aligned} \tag{22}$$

As suggested by Fig. 2, the ratio $\sigma_{\widetilde{\pi}}/\sigma_{\pi^\star}$ is larger than 1 for any γ and tends to 1 when $\gamma \to 1$.

In our original postgrowth experiments [2], the sub-Poissonian statistics of nucleation were demonstrated and quantified by counting the numbers of nucleations occurring in fixed time intervals. The full calculation of the probabilities for any number of nucleations in any fixed time interval will be carried out in Sec. IV. There is however another quantity that should display a divergence from Poisson statistics, namely the distribution of waiting times between nucleations. This cannot be extracted from our postgrowth observations. On the other hand, the *in situ* TEM experiments, which usually proceed at fairly low growth rates [13–15], can detect individual nucleation events and locate them accurately in time. Nothing thus opposes the systematic study of the distribution of waiting times, although no such study seems to exist yet. In Sec. V, we calculate this distribution analytically.

IV. PROBABILITY FOR A GIVEN NUMBER OF NUCLEATIONS IN A TIME INTERVAL

The sub-Poissonian nucleation statistics in VLS growth were first demonstrated by studying the distribution of the number of nucleations occurring in fixed time intervals [2]. In this section, we calculate this quantity analytically in the framework of our model.

A. Conditional probability

We consider arbitrary time intervals of fixed length T. We first assume that the NP at the start of the interval is P_0 and define $\pi_n(T|P_0)$ as the probability that exactly n nucleations ($n \geqslant 0$) will occur during time T.

The probability for having no nucleation over the entire interval is

$$\pi_0(T|P_0) = \exp\left(-\int_0^T P_0 e^{Ab\tau} d\tau\right) = \exp(-\beta E P_0), \tag{23}$$

with

$$E(T) = e^{AbT} - 1 = e^{T/\beta} - 1. \tag{24}$$

Since $\bar{\tau} = -\beta\ln\gamma$, E depends on T only via $T/\bar{\tau}$. In the following, whenever no confusion is possible, the value of E for the time currently considered is simply noted E.

The probability for having exactly one nucleation over the same interval is

$$\pi_1(T|P_0) = \int_0^T d\tau\, \pi_0(\tau|P_0)\, P_0 e^{Ab\tau}\pi_0(T-\tau|P_0 e^{A(b\tau-\delta)}), \tag{25}$$

where, in the integral, the first π_0 factor accounts for the probability of having no nucleation before time τ ($0 \leqslant \tau \leqslant T$) has elapsed, $P_0 e^{Ab\tau} d\tau$ for the probability of having a nucleation at time τ (within $d\tau$), at which the NP is $P_0 e^{Ab\tau}$, and the second π_0 factor for the probability of having again no nucleation after this, taking into account the abrupt decrease of the NP induced by nucleation at time τ. Using Eq. (23), we find

$$\pi_1(T|P_0) = \frac{1}{1-\gamma}[\exp(-\gamma\beta E P_0) - \exp(-\beta E P_0)]. \tag{26}$$

Similarly,

$$\pi_2(T|P_0) = \int_0^T d\tau\, \pi_0(\tau|P_0)\, P_0 e^{Ab\tau}\pi_1(T-\tau|P_0 e^{A(b\tau-\delta)}). \tag{27}$$

In this integral, the first two factors have the same meaning as in the calculation of π_1. The π_1 factor accounts for the probability of a second nucleation (and only one) occurring in the remaining time $T-\tau$, taking into account that the NP is $P_0 e^{Ab\tau}$ immediately before the first nucleation and hence $P_0 e^{A(b\tau-\delta)}$ immediately after. Using Eqs. (23) and (26), we

get

$$\pi_2(T|P_0) = \frac{1}{(1-\gamma)^2}\left[\frac{\gamma}{1+\gamma}\exp(-\beta E P_0) - \exp(-\gamma\beta E P_0) + \frac{1}{1+\gamma}\exp(-\gamma^2\beta E P_0)\right]. \quad (28)$$

Examination of the expressions of π_0, π_1, and π_2 leads us to propose the general formula

$$\pi_n(T|P_0) = \frac{1}{(1-\gamma)^n}\sum_{p=0}^{n} a_p^{(n)} \exp(-\gamma^p \beta E P_0). \quad (29)$$

In the Appendix (Sec. 7), we demonstrate this formula and derive the values of the coefficients $a_p^{(n)}$ [Eq. (A32)], which readily yield

$$\pi_n(T|P_0) = \sum_{p=0}^{n} \frac{g_{n-p}}{(\gamma;\gamma)_p} \exp(-\gamma^p \beta E P_0). \quad (30)$$

B. Total probability

We now use the results of Secs. III and IV A to calculate the probability $\pi_n(T)$ that exactly n nucleations occur over an *arbitrary* time interval of duration T. Arbitrary means in particular that we ignore if there is a nucleation at the beginning of the interval or not. Since we defined $\pi^\star$ as the density of probability of the NP over the whole process (Sec. III A), arbitrarily choosing the beginning of the interval is equivalent to having a probability $\pi^\star(P)dP$ for the NP to be equal to P, within dP, at this instant. Hence

$$\pi_n(T) = \int_0^\infty \pi_n(T|P)\pi^\star(P)dP. \quad (31)$$

Using Eqs. (30) and (17), this gives

$$\pi_n(T) = -\frac{1}{(\gamma;\gamma)_\infty \ln\gamma}\sum_{r=0}^{n}\frac{g_{n-r}}{(\gamma;\gamma)_r}\int_0^\infty \frac{dP}{P} \times \sum_{m=0}^{\infty} g_m \exp[-(\gamma^r\beta E + \gamma^{-m}\beta)P]. \quad (32)$$

Using Eq. (A19) with $F_m = 1+\gamma^{r+m}E$ to calculate the integral (at given r), we obtain

$$\pi_n(T) = \frac{1}{(\gamma;\gamma)_\infty \ln\gamma}\sum_{r=0}^{n}\frac{g_{n-r}}{(\gamma;\gamma)_r} \times \left[\sum_{m=0}^{\infty} g_m \ln(1+\gamma^{r+m}E) + (\gamma;\gamma)_\infty \ln\gamma\right] = \delta_{n0} + \frac{1}{(\gamma;\gamma)_\infty \ln\gamma}\sum_{r=0}^{n}\frac{g_{n-r}}{(\gamma;\gamma)_r} \times \sum_{m=0}^{\infty} g_m \ln(1+\gamma^{r+m}E), \quad (33)$$

where the second equality makes use of Eq. (A10) of the Appendix, Sec. 3. Since $E(T=0)=0$, we immediately check that this formula gives the expected results for $T=0$, namely unit probability for no nucleation and zero probability for any nonzero number of nucleations. Since E depends on T only via $T/\bar{\tau}$ [Eq. (24)], the same holds for $\pi_n(T)$, as expected. Equation (33) is not summable in closed form. However, if γ is not too close to 1, the infinite sum converges rapidly and excellent approximations are obtained by considering only a small number of its terms.

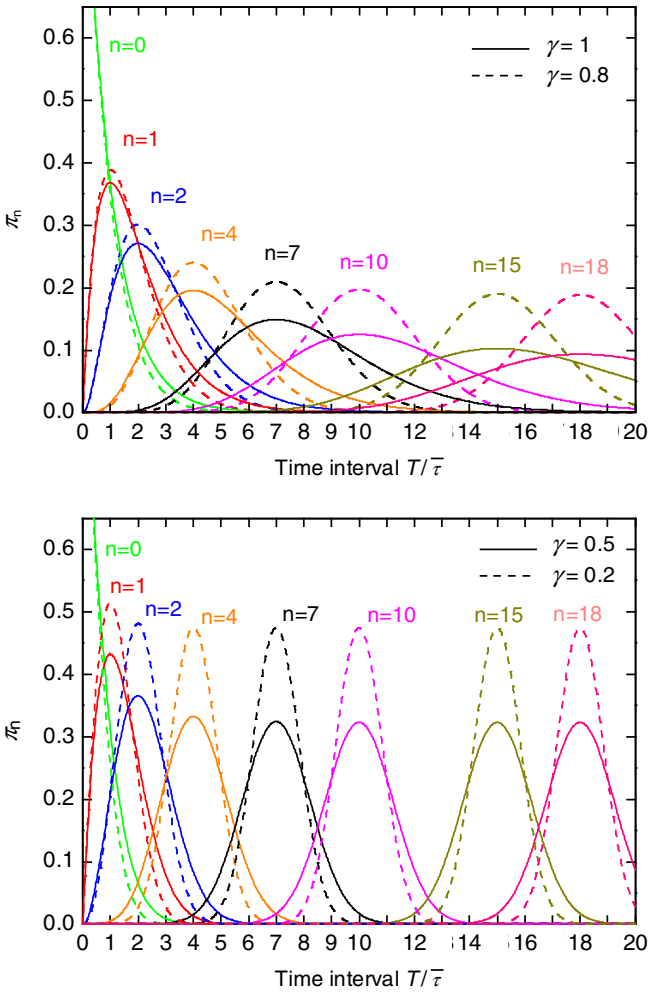

FIG. 3. (Color online) Variation with length of time interval T (in units of average waiting time $\bar{\tau}$) of the probabilities π_n for having given numbers n of nucleation during this interval, for the Poisson process ($\gamma = 1$; top panel) and moderately ($\gamma = 0.8$; top panel) to strongly ($\gamma = 0.5$ and $\gamma = 0.2$; bottom panel) sub-Poissonian processes.

These calculations are illustrated in Figs. 3 and 4 for a broad range of values of γ. Namely, Fig. 3 shows the variations of π_n with time interval T for a set of values of n and Fig. 4 the distribution of the various π_n for selected values of T. As expected, as soon as $\gamma < 1$, this distribution becomes narrower than the Poisson distribution and all the more so that γ is small, in perfect agreement with our previous findings [2]. This narrowing will be fully quantified in Sec. IV C 2 and further discussed in Sec. VI. In addition, a shift and change of symmetry of the distribution of π_n at fixed T also appears from Fig. 4. It is most easily seen for time intervals T_m equal to an integer number m of average waiting times $\bar{\tau}$: whereas in the Poisson case, the π_n are equal and maximum for $n = m-1$ and $n = m$, for even moderately sub-Poissonian processes, π_n is largest for $n = m$, with $\pi_{m-1} \simeq \pi_{m+1}$, unless m is very small.

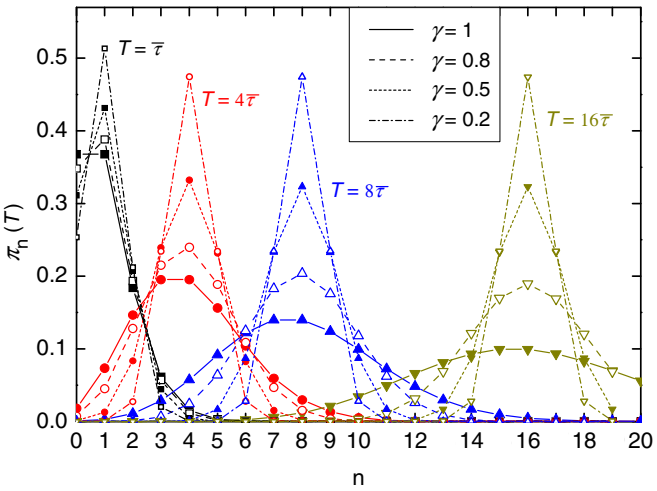

FIG. 4. (Color online) Distribution of the probabilities $\pi_n(T)$ for time intervals T equal to 1 (squares), 4 (circles), 8 (up triangles), and 16 (down triangles) average waiting times $\bar{\tau}$, for the Poisson process ($\gamma = 1$; large full symbols, full lines) and various sub-Poissonian processes: $\gamma = 0.8$ (large empty symbols, dashed lines), $\gamma = 0.5$ (small full symbols, short dashes), and $\gamma = 0.2$ (small empty symbols, dash-dotted lines). Since the probabilities are defined solely for integer values, the lines are only guides to the eye.

C. Statistical properties

In this section, we calculate the average $\langle n(T)\rangle$ and the standard deviation $\sigma(T)$ of the number $n(T)$ of nucleations occurring over a fixed time T. Of course, the average can be straightforwardly obtained by noting that, since refilling provides bT atoms, in the stationary regime, one must have $\langle n(T)\rangle = bT/\delta$. However, in Sec. IV C 1, we will derive this result by using only the set of probabilities π_n, as a check of our calculation [Eq. (33)], but also to introduce and compute several quantities of use to evaluate $\sigma(T)$ in Sec. IV C 2.

1. Average

The average value $\langle n(T)\rangle$ of the number of nucleations $n(T)$ during time T is

$$\langle n(T)\rangle = \sum_{p=0}^{\infty} p\pi_p(T). \tag{34}$$

Hence $\langle n(T)\rangle = \frac{\partial F}{\partial z}|_{z=1}$, where F is the appropriate generating function:

$$F(T,z) = \sum_{n=0}^{\infty} \pi_n(T) z^n. \tag{35}$$

To calculate $\partial F/\partial z$, we will first compute $\partial F/\partial E$. Equation (33) yields

$$\begin{aligned}\frac{\partial \pi_n}{\partial E} &= \frac{1}{(\gamma;\gamma)_\infty \ln\gamma} \sum_{p=0}^{n} \frac{g_{n-p}\gamma^p}{(\gamma;\gamma)_p} S(\gamma^p E) \\ &= \frac{1}{(-E;\gamma)_\infty \ln\gamma} \sum_{p=0}^{n} \frac{g_{n-p}\gamma^p(-E;\gamma)_p}{(\gamma;\gamma)_p}.\end{aligned} \tag{36}$$

Here, $S(x)$ is defined by

$$S(x) = \sum_{n=0}^{\infty} \frac{g_n}{x+\gamma^{-n}} \tag{37}$$

and, to establish the second equality, we used the calculation of S carried out in Sec. 4 of the Appendix [Eq. (A15)] and noted that $(-\gamma^p E;\gamma)_\infty = (-E;\gamma)_\infty/(-E;\gamma)_p$. Hence

$$\begin{aligned}\frac{\partial F}{\partial E} &= \frac{1}{(-E;\gamma)_\infty \ln\gamma} \sum_{n=0}^{\infty} z^n \sum_{p=0}^{n} \frac{g_{n-p}\gamma^p(-E;\gamma)_p}{(\gamma;\gamma)_p} \\ &= \frac{1}{(-E;\gamma)_\infty \ln\gamma} \sum_{p=0}^{\infty} \frac{\gamma^p(-E;\gamma)_p}{(\gamma;\gamma)_p} \sum_{n=p}^{\infty} g_{n-p} z^n \\ &= \frac{(z;\gamma)_\infty}{(-E;\gamma)_\infty \ln\gamma} \sum_{p=0}^{\infty} \frac{(\gamma z)^p(-E;\gamma)_p}{(\gamma;\gamma)_p},\end{aligned} \tag{38}$$

where the second equality was obtained by reordering the terms of the double sum and the third one by using Eq. (10). Provided $|\gamma z| < 1$, the last sum exists and is given by the (infinite) q-binomial theorem [29], so that

$$\frac{\partial F}{\partial E} = \frac{(z;\gamma)_\infty}{(-E;\gamma)_\infty \ln\gamma} \frac{(-E\gamma z;\gamma)_\infty}{(\gamma z;\gamma)_\infty} = \frac{(1-z)(-E\gamma z;\gamma)_\infty}{(-E;\gamma)_\infty \ln\gamma}. \tag{39}$$

Hence

$$\begin{aligned}\frac{\partial^2 F}{\partial z \partial E} = \frac{1}{(-E;\gamma)_\infty \ln\gamma}\Bigg[&-(-E\gamma z;\gamma)_\infty + (1-z) \\ &\times \sum_{p=1}^{\infty} \frac{E\gamma^p(-E\gamma z;\gamma)_\infty}{1+E\gamma^p z}\Bigg].\end{aligned} \tag{40}$$

In Eq. (40), the series is the z derivative of the infinite product $(-E\gamma z;\gamma)_\infty$. This series converges at least for any real positive value of z, since the series of generic term $\gamma^p/(1+E\gamma^p z)$ itself converges. Hence

$$\frac{\partial}{\partial E}\left(\left.\frac{\partial F}{\partial z}\right|_{z=1}\right) = \frac{-(-E\gamma;\gamma)_\infty}{(-E;\gamma)_\infty \ln\gamma} = -\frac{1}{(1+E)\ln\gamma}. \tag{41}$$

After integration, noting that $\langle n(T=0)\rangle = 0$ (since $\pi_n = \delta_{n0}$ when $T = 0$),

$$\left.\frac{\partial F}{\partial z}\right|_{z=1}(E) = -\frac{\ln(1+E)}{\ln\gamma}. \tag{42}$$

Recalling that this quantity equals $\langle n(T)\rangle$ and returning to the definitions of γ and E [Eq. (24)], we finally obtain the expected result, namely

$$\langle n(T)\rangle = \frac{bT}{\delta} = \frac{T}{\bar{\tau}}. \tag{43}$$

2. Standard deviation

The standard deviation σ of the number of nucleations occurring during time T is

$$\begin{aligned}\sigma^2(T) &= \langle n^2(T)\rangle - \langle n(T)\rangle^2 \\ &= \left.\frac{\partial^2 F}{\partial z^2}\right|_{z=1} + \left.\frac{\partial F}{\partial z}\right|_{z=1} - \left(\left.\frac{\partial F}{\partial z}\right|_{z=1}\right)^2.\end{aligned} \tag{44}$$

From Eq. (40),

$$\frac{\partial^3 F}{\partial z^2 \partial E} = \frac{1}{(-E;\gamma)_\infty \ln\gamma}\left\{ -2\sum_{p=1}^{\infty}\frac{E\gamma^p(-E\gamma z;\gamma)_\infty}{1+E\gamma^p z} + (1-z)(-E\gamma z;\gamma)_\infty\left[\left(\sum_{p=1}^{\infty}\frac{E\gamma^p}{1+E\gamma^p z}\right)^2 - \sum_{p=1}^{\infty}\frac{(E\gamma^p)^2}{(1+E\gamma^p z)^2}\right]\right\}. \quad (45)$$

As in the case of the average, the three series appearing in Eq. (45) converge. Then

$$\frac{\partial}{\partial E}\left(\left.\frac{\partial^2 F}{\partial z^2}\right|_{z=1}\right) = \frac{-2(-E\gamma;\gamma)_\infty}{(-E;\gamma)_\infty \ln\gamma}\sum_{p=1}^{\infty}\frac{E\gamma^p}{1+E\gamma^p} = -\frac{2}{\ln\gamma}\frac{U(E)}{1+E}, \quad (46)$$

with

$$U(E) = \sum_{p=1}^{\infty}\frac{E\gamma^p}{1+E\gamma^p}. \quad (47)$$

From Eq. (44), and then Eqs. (41), (42), and (46), we find

$$\frac{\partial\sigma^2}{\partial E} = \frac{\partial}{\partial E}\left(\left.\frac{\partial^2 F}{\partial z^2}\right|_{z=1}\right) + \frac{\partial}{\partial E}\left(\left.\frac{\partial F}{\partial z}\right|_{z=1}\right) - 2\left.\frac{\partial F}{\partial z}\right|_{z=1}\frac{\partial}{\partial E}\left(\left.\frac{\partial F}{\partial z}\right|_{z=1}\right) = -\frac{1}{\ln\gamma}\left[\frac{1}{1+E} + \frac{2}{\ln\gamma}\frac{\ln(1+E)}{1+E} + 2\frac{U(E)}{1+E}\right]. \quad (48)$$

For $T = 0$, $E = 0$, and, as already noted, $\pi_n = \delta_{n0}$, so that $\sigma^2 = 0$. Integrating Eq. (48), using Eq. (47), thus gives, with $E = E(T)$,

$$\sigma^2(E) = -\frac{1}{\ln\gamma}\left\{\ln(1+E) + \frac{1}{\ln\gamma}[\ln(1+E)]^2 + 2\sum_{p=1}^{\infty}\frac{1}{1-\gamma^p}\ln\frac{1+E\gamma^p}{(1+E)^{\gamma^p}}\right\}. \quad (49)$$

Equation (49) is convenient for estimating $\sigma(E)$ since, for most values of γ, it suffices to take into account a small number of terms in the series. However, it is not adapted to discussing the behavior of $\sigma(E)$ for $E \to \infty$. To this end, we return to Eq. (47) and compute the series using the Euler-Maclaurin formula [32,33]. We may indeed rewrite

$$\frac{U(E)}{E} = \sum_{p=0}^{\infty} f_E(p) - \frac{1}{1+E}, \quad (50)$$

where, for a given γ,

$$f_E(x) = \frac{\gamma^x}{1+E\gamma^x}. \quad (51)$$

The derivatives of function f_E are calculated in the Appendix, Sec. 6 [Eq. (A21)]. Since f_E and all its derivatives tend to zero when $x \to \infty$, Euler-Maclaurin's formula at order $k \geqslant 1$ is written as

$$\frac{U(E)}{E} = \frac{-1}{1+E} + \int_0^\infty f_E(t)dt - \sum_{i=1}^{k}\frac{b_i}{i!}f_E^{(i-1)}(0) + (-1)^{k+1}\int_0^\infty \frac{B_k(\{t\})}{k!}f_E^{(k)}(t)dt, \quad (52)$$

where b_n and B_n are, respectively, the Bernoulli number and Bernoulli polynomial of rank n and where $\{x\}$ designates the fractional part of x [32,33]. Recall that all Bernoulli numbers of odd rank are zero, except $b_1 = -1/2$. Equation (52), where the discrete sum only appears if $k \geqslant 2$, is valid provided the last integral converges. In this case, substituting Eq. (52) into Eq. (48) after calculating the first integral, we get, by integration over E and using the expressions of the derivatives of f_E [Eq. (A21)],

$$\sigma^2(E) = -\frac{1}{\ln\gamma}\frac{E}{1+E} + 2\sum_{i=2}^{k}\frac{b_i}{i!}(\ln\gamma)^{i-2} \times \int_0^E \frac{\epsilon A_{i-1}(-\epsilon)}{(1+\epsilon)^{i+1}}d\epsilon + R_k(E). \quad (53)$$

In Eq. (53), A_n is the nth Eulerian polynomial [34] and the remainder R_k is as

$$R_k(E) = (-1)^k\frac{2}{\ln\gamma}\int_0^\infty dt\frac{B_k(\{t\})}{k!}\int_0^E\frac{\epsilon}{1+\epsilon}f_\epsilon^{(k)}(t)d\epsilon. \quad (54)$$

Setting $u = \gamma^t$, the remainder is rewritten as

$$R_k(E) = 2\frac{(-1)^{k+1}}{k!}(\ln\gamma)^{k-2}\int_0^\infty du\, B_k\left(\left\{\frac{\ln u}{\ln\gamma}\right\}\right) \times \int_0^E \frac{\epsilon A_k(-\epsilon u)}{(1+\epsilon)(1+\epsilon u)^{k+1}}d\epsilon. \quad (55)$$

In particular, for $k = 2$, recalling that $b_2 = 1/6$ and $A_2(x) = 1 + x$, we find

$$\sigma^2(E) = -\frac{1}{\ln\gamma}\frac{E}{1+E} + \frac{1}{12}\left(\frac{E}{1+E}\right)^2 + R_2(E), \quad (56)$$

with

$$R_2(E) = -\int_0^1 B_2\left(\left\{\frac{\ln u}{\ln\gamma}\right\}\right)V_{E,2}(u)du \quad (57)$$

and

$$V_{E,2}(u) = \frac{E(u+1)+2E^2u^2}{(u-1)^2(1+Eu)^2} + \frac{u+1}{(u-1)^3}\ln\frac{1+E}{1+Eu}. \quad (58)$$

In Eq. (57), the fractional part varies between zero and 1. Since the Bernoulli polynomials are bounded over this interval and since, for $u \to 1^-$, $V_{E,2}(u) = \frac{(3-E)E^2}{6(1+E)^3} + O(1-u)$, the integral in Eq. (57) converges for any finite E.

We demonstrate in the Appendix (Sec. 8) that, when $\gamma < 1$, σ^2 remains finite when $E \to \infty$. This is of course in sharp contrast with the infinite increase of σ^2 in the Poissonian case ($\gamma = 1$, $\sigma^2 = T/\bar{\tau}$), but in agreement with one of the major results of our previous combined experimental and

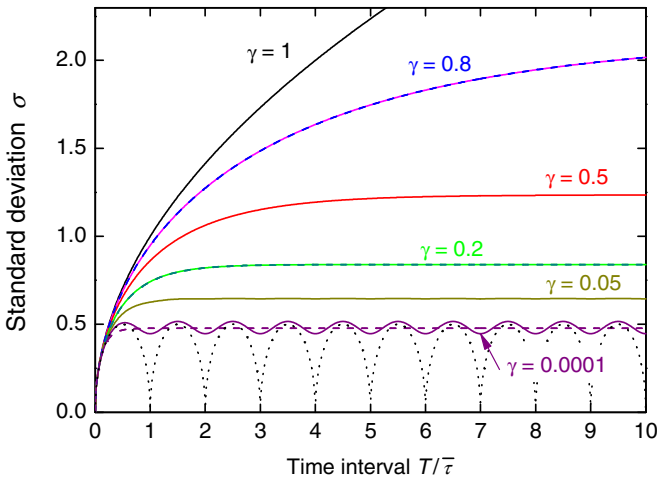

FIG. 5. (Color online) Full curves: variation with width of time interval T of the standard deviation σ of the number of nucleation events occurring during T, for the Poisson process ($\gamma = 1$) and for moderately ($\gamma = 0.8$) to extremely ($\gamma = 0.0001$) sub-Poissonian processes. The time unit is set to $\bar{\tau}$, the average time between nucleations. Dashed curves: standard deviations approximated by omitting remainder R_2 in Eq. (56), for $\gamma = 0.8$, 0.2, and 0.0001. Dotted curve: limit of the standard deviation for a periodic process ($\gamma = 0$).

numerical study [2]. This is illustrated in Fig. 5, which shows the variations of σ with time interval T for selected values of γ. As a complement, Fig. 6 gives the variations of σ with γ for a range of time intervals.

Equation (56), together with Eqs. (57) and (58), gives the exact value of σ for any time interval T. The numerical evaluation of the remainder [Eq. (57)] is easily carried out. However, unless γ is very small, Eq. (56) with remainder R_2 omitted gives an excellent approximation of σ^2: as illustrated in Fig. 5 (dashes), this holds not only for moderately ($\gamma = 0.8$)

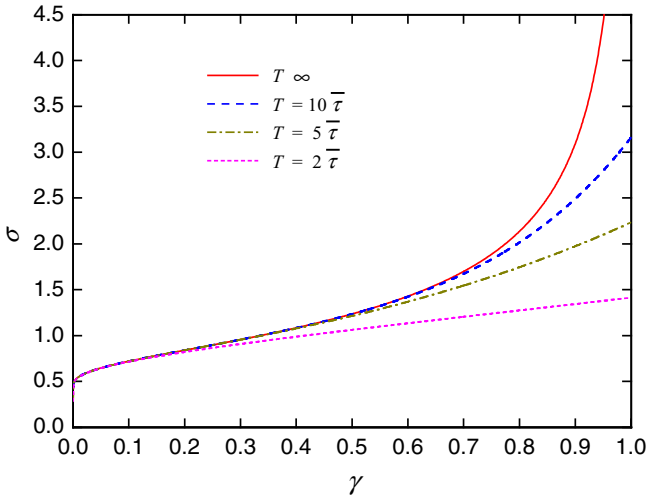

FIG. 6. (Color online) Variation with parameter γ of the standard deviation σ of the number of nucleation events occurring during T, for T equal to $2\bar{\tau}$ (short dashes), $5\bar{\tau}$ (dash-dotted curve), or $10\bar{\tau}$ (dashes) and for T infinite (full curve).

but also for strongly ($\gamma = 0.2$) sub-Poissonian processes. It is only for $\gamma \ll 1$ that R_2 adopts an oscillatory behavior with a period equal to $\bar{\tau}$. In this case, Eq. (56) with R_2 omitted still gives an excellent approximation of the average value of σ over any interval of width $\bar{\tau}$ (case $\gamma = 0.0001$ in Fig. 5). In the limit $\gamma \to 0$, nucleation becomes periodic and $\sigma^2(T)$ tends to the corresponding value, namely $\{\frac{T}{\bar{\tau}}\}(1 - \{\frac{T}{\bar{\tau}}\})$ (dots in Fig. 5). The behavior of σ will be further discussed in Sec. VI.

V. WAITING TIMES

The density of probability of the waiting time τ between two consecutive nucleations, per unit waiting time, is simply

$$\hat{\pi}(\tau) = \int_0^\infty \pi(\tau|P_0)\widetilde{\pi}(P_0)dP_0, \tag{59}$$

with $\pi(\tau|P_0)$ the conditional probability and $\widetilde{\pi}$ the density of probability of nucleation given respectively by Eqs. (3) and (4) and Eq. (13). Hence

$$\hat{\pi}(\tau) = \gamma\, e^{\tau/\beta} \int_0^\infty P_0\, e^{-\gamma\beta E(\tau)P_0}\widetilde{\pi}(P_0)dP_0, \tag{60}$$

with $E(\tau)$ defined by Eq. (24). The integration is readily performed and yields

$$\hat{\pi}(\tau) = \frac{E(\tau)+1}{\beta(\gamma;\gamma)_\infty}\,\Phi(E(\tau)), \tag{61}$$

with

$$\Phi(E) = \gamma \sum_{n=0}^{\infty} \frac{g_n}{(\gamma^{-n}+\gamma E)^2}. \tag{62}$$

Density $\hat{\pi}$ can easily be calculated using this formula since, in most cases, only the first few terms contribute effectively to the sum. Alternatively, we may rewrite

$$\Phi(E) = -\frac{\partial}{\partial E}[S(\gamma E)]. \tag{63}$$

Here, $S(x)$ is defined by Eq. (37) so that, from Eq. (A15), $\ln S(x) = \ln(\gamma;\gamma)_\infty - \sum_{p=0}^{\infty}\ln(1+x\gamma^p)$ and

$$\frac{dS}{dx} = -S(x)\left[\frac{U(x)}{x} + \frac{1}{1+x}\right]. \tag{64}$$

Euler-Maclaurin's development of function U [Eq. (52)] at order 2 yields

$$\hat{\pi}(\tau) = \frac{1}{\beta(\gamma;\gamma)_\infty}\frac{E+1}{E}S(\gamma E)\left[-\frac{\ln(1+\gamma E)}{\ln\gamma} + \frac{\gamma E}{2(1+\gamma E)} - \frac{\gamma(\ln\gamma)E}{12(1+\gamma E)^2} + \frac{\gamma(\ln\gamma)E}{2}I_2(E)\right], \tag{65}$$

where E stands for $E(\tau)$ and

$$I_2(E) = \int_0^1 B_2\left(\left\{\frac{\ln u}{\ln\gamma}\right\}\right)\frac{1-\gamma E u}{(1+\gamma E u)^3}du. \tag{66}$$

The numerical evaluation of integral I_2 provides another efficient means of calculating $\hat{\pi}(\tau)$. Figure 7 illustrates the variations of $\hat{\pi}$ with T for a broad range of values of γ.

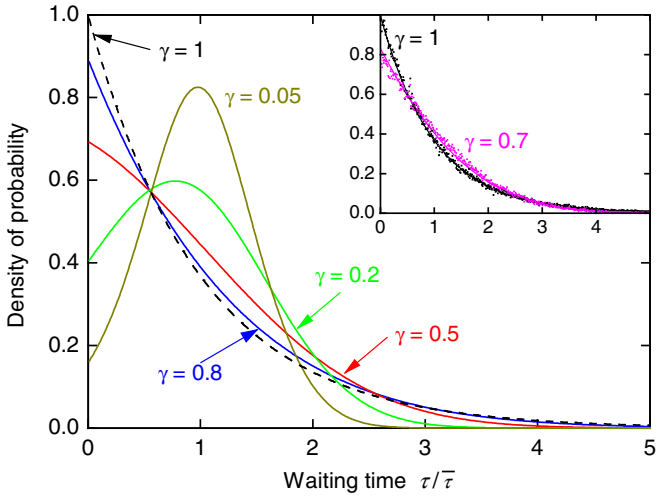

FIG. 7. (Color online) Main panel: probability distribution of the waiting times calculated for the Poisson process ($\gamma = 1$; dashes) and for moderately ($\gamma = 0.8$) to very strongly ($\gamma = 0.05$) sub-Poissonian processes (full curves). The time unit is set to $\bar{\tau}$, the average time between nucleations. Inset: comparison of the distribution of waiting times calculated according to the present model (lines) with the simulated distribution (symbols), for $\gamma = 1$ and $\gamma = 0.7$.

VI. DISCUSSION AND CONCLUSIONS

In the preceding sections, we gave a comprehensive treatment of the statistics of nucleation in the sub-Poissonian regime, based on a simple model that describes the cyclic *albeit aperiodic* depletion of a supersaturated parent phase of nanoscopic dimensions upon nucleation followed by the rapid growth of a small crystal building block (such as a nanowire monolayer). The sequence of nucleation events is modeled as a stochastic Markov process. We assume that nucleation is governed by the variations of the concentration of a single minority constituent. We previously proposed this model to account for our experimental observation of a self-regulation of the nucleation events in VLS-grown III–V NWs and showed by numerical simulations that it could reproduce quantitatively the measured statistics, given a proper choice of the model parameters [2]. Since the present calculations assume that $\Delta\mu$ and hence P depend only on N [Eq. (1)], they cannot be used if a mixture of crystalline structures appears: because of differences of cohesive energy, the probability of nucleation then also depends on the structure of the phase formed. Such a sphalerite/wurtzite polytypism may appear in NWs of compound semiconductors and the probability of nucleation of a new ML is indeed known to depend on the nature (cubic or hexagonal) of this ML [12]. On the other hand, various single phase crystals may be considered by simply taking into account the slight change of constant A induced by the shift of chemical potential [2]. Note that our experimental study [2] was performed on such single phase NWs.

In the present work, the main assumptions are first that the consumption of material to build the finite growth unit happens instantaneously after nucleation and, second, that the chemical potential of the parent phase varies linearly with the concentration of the species considered and hence that the nucleation probability depends exponentially on it. The latter was shown to be a good hypothesis for our NWs [2] (namely, simulations carried out with or without this hypothesis match closely). In these conditions, the deviation from Poisson statistics is quantified by a single parameter γ, the ratio of the nucleation probabilities immediately after and before nucleation (hence $0 \leqslant \gamma \leqslant 1$ and the more sub-Poissonian the process, the smaller γ). This allows us to carry out fully analytical calculations using q-calculus. These calculations follow from a self-consistent determination of two densities of probability of the nucleation probability P. The first density, $\widetilde{\pi}$, is restricted to those instants when nucleation occurs, the second one, $\pi^\star$, covers the whole process. We showed in particular that the former is a known solution of the so-called pantograph differential equation. From these densities, we calculated other statistical quantities amenable to comparison with experiments and simulations, in particular the distribution of the probabilities $\pi_n(T)$ of having n nucleations occurring in an arbitrary time interval of given length T [Sec. IV B; Eq. (33)], its standard deviation [Sec. IV C 2; Eq. (49) or (56)] and the distribution of the waiting times between nucleations [Sec. V; Eq. (65)]. For all these quantities, the agreement between the present analytical calculations and our simulations is excellent, as illustrated for the waiting times in the inset of Fig. 7.

The statistical distributions of yet more quantities may be derived from our explicit determination of densities $\pi^\star(P)$ and $\widetilde{\pi}(P)$. Indeed, our model postulates a one to one correspondence between nucleation probability P and number N of atoms in the parent phase [Eq. (1)]. Hence, the distribution of probability of N over the whole process is straightforwardly obtained from $\pi^\star$ and the distribution of this number, restricted to those instants where nucleation occurs, from $\widetilde{\pi}$. Since there is also a one to one relation between N and $\Delta\mu$, the same holds for the statistical distribution of $\Delta\mu$, but this requires selecting a particular system for which the constants specifying this relation, which are not primary parameters of our model, may be obtained. Similarly, in the case of two-dimensional nuclei (as usually pertains for NWs), the radius of the critical nucleus scales inversely with $\Delta\mu$; hence the distribution of the radii of these nuclei also follows straightforwardly from $\widetilde{\pi}$, but again the inverse proportionality constant is not a primary model parameter.

To our knowledge, in addition to our initial work [2], three studies so far have tackled the calculation of limited aspects of the nucleation statistics. Kovalchuk *et al.* consider point contact growth of NWs and concentrate on the distribution of the waiting times [9], whereas Sibirev evaluates the distribution of the number of nucleations in a given time interval [35] and Dubrovskii the evolution of the NP [17] in VLS NW growth. One of the major differences with the present work is that these authors all restrict themselves to the relatively simple problem of the evolution of the system after some initial time at which its state is assumed to be known (Kovalchuk *et al.* state that they "neglect the time correlation between subsequent steps" [9], Sibirev fixes "the amount of material in the drop. . .at the initial instant" [35], and Dubrovskii "assumes that P equals zero at $t = 0$" [17]). In other words, they only calculate the NP *conditional* to a given state of the system at a given time (the start of the waiting period [9] or the start of the interval over which nucleations are counted [35]). Although

we also provide analytical expressions for these conditional probabilities $\pi(P|P_0)$ and $\pi(\tau|P_0)$ [Eqs. (3) and (4)], they are primarily steps toward the calculation of distributions that take into account *self-consistently* the probabilities of all possible initial states. It is only in the limits of long waiting times [9] or long counting intervals [35] that the two approaches converge, the memory of the initial state of the system being nearly forgotten. This is manifested for instance by the fact that imposing initial conditions misses the increase of the maximum of $\pi_n(T)$ as n decreases [compare our Fig. 3 with Fig. 1(b) of Ref. [35]].

Perhaps the most striking manifestation of the effect studied here is the narrowing of the distribution of the numbers $\pi_n(T)$ of nucleations occurring over time intervals of given length T. This was demonstrated experimentally and theoretically in our initial work [2] and is fully confirmed by the present calculations. Figure 4 shows that, even for moderately sub-Poissonian processes, the distribution is much narrower than the Poisson distribution and all the more so when the time interval is long. This figure also makes it clear that, at variance with the Poisson case, the width of the π_n distribution hardly increases with T (unless this interval is very short), which is confirmed quantitatively by Sec. IV C 2 and illustrated in Fig. 5.

It is then tempting to imagine that the sub-Poissonian process is quasiperiodic, with nucleation events occurring at regular intervals dictated by the time needed to replace in the nanosized mother phase the atoms consumed in forming the relevant growth unit after nucleation. The idea behind this is that nucleation is very unlikely if supersaturation is too low but becomes very likely when it reaches some critical value [25], so that all nucleations will occur at more or less the same value of supersaturation and hence at fairly regular intervals [17]. However, our calculations show that this is not the case, at least for values of the parameters that pertain to the systems currently studied in this respect (recall that, for our Au-catalyzed or self-catalyzed NWs, we estimated that γ varies between 0.06 and 0.7). This is most apparent from the distribution of the waiting times (Fig. 7). A well-known feature of the Poisson process ($\gamma = 1$) is that the waiting times are exponentially distributed; in particular, the most probable waiting time is zero. Figure 7 shows that the monotonously decreasing character of the distribution is preserved over a wide range of γ values. It is only for $\gamma < \gamma_c = \sqrt{2} - 1 \simeq 0.414$ that the distribution presents a maximum at a certain nonzero waiting time. And even for much lower values of γ, the distribution remains very broad with a significant probability of having waiting times very short compared to the average time between nucleations, or on the other hand much longer (this is also apparent from the simulation of Fig. 1). This contrasts sharply with quasiperiodic nucleation, which would produce a narrow peak. Strikingly, for our In(As,P) NWs [2] that display a strong reduction of σ, the calculated distribution of waiting times is only subtly modified with respect to Poisson statistics (Fig. 7, inset), with slightly less short and long waiting times and slightly more in the middle range. This proves that the narrowing of the π_n distribution does not imply at all that nucleation can only happen in a narrow time window; actually, nucleation can happen at any time, even though its probability is seriously reduced after a first nucleation.

In summary, our comprehensive calculations provide analytical expressions that describe quantitatively a wide range of statistical properties of the processes of self-regulated sub-Poissonian nucleation in a nanophase. This already allowed us to discuss various manifestations of this effect and to clarify some apparent contradictions between these manifestations. The way is now open for further predictions and comparisons with experiments or simulations. Our calculations also provide useful guides to set conditions where these effects can be put to profit to limit the randomness inherent in nucleation and therefore to increase the uniformity of nanostructures.

APPENDIX

1. Calculation of series $S^{(0)}$

We first calculate the following series:

$$S^{(0)} = \sum_{n=0}^{\infty} n g_n. \tag{A1}$$

We note that $S^{(0)} = \frac{dG}{dz}(z=1)$, where

$$G(z) = \sum_{n=0}^{\infty} g_n z^n. \tag{A2}$$

From Eq. (10), $G(z) = (z;\gamma)_\infty$, and thus, noting that $(z;\gamma)_\infty = (1-z)(\gamma z;\gamma)_\infty$,

$$\frac{dG}{dz} = \frac{d}{dz}\prod_{p=0}^{\infty}(1-\gamma^p z) = \sum_{p=0}^{\infty} -\gamma^p \frac{(z;\gamma)_\infty}{1-\gamma^p z}$$
$$= -(\gamma z;\gamma)_\infty \left\{ 1 + (1-z)\sum_{p=1}^{\infty} \frac{\gamma^p}{1-\gamma^p z} \right\}. \tag{A3}$$

Therefore,

$$S^{(0)} = -(\gamma;\gamma)_\infty. \tag{A4}$$

2. Calculation of sums $S_n^{(1)}$

Here, we show that sum $S_n^{(1)}$ defined by

$$S_n^{(1)} = \sum_{p=0}^{n} (-1)^p \gamma^{\frac{(p+1)(p-2n)}{2}} \frac{(\gamma;\gamma)_{n+1}}{(\gamma;\gamma)_{p+1}(\gamma;\gamma)_{n-p}} \tag{A5}$$

is uniformly equal to 1. To simplify the notation, we introduce the q-binomial coefficients [29]:

$$\binom{n}{p}_\gamma = \frac{(\gamma;\gamma)_n}{(\gamma;\gamma)_p(\gamma;\gamma)_{n-p}} \quad \text{for } 0 \leqslant p \leqslant n. \tag{A6}$$

Then

$$S_n^{(1)} = \sum_{p=0}^{n} (-1)^p \gamma^{\frac{(p+1)(p-2n)}{2}} \binom{n+1}{p+1}_\gamma$$
$$= -\sum_{m=1}^{n+1} (-\gamma^{-n})^m \gamma^{\frac{m(m-1)}{2}} \binom{n+1}{m}_\gamma$$
$$= 1 - \sum_{m=0}^{n+1} (-\gamma^{-n})^m \gamma^{\frac{m(m-1)}{2}} \binom{n+1}{m}_\gamma$$
$$= 1 - (\gamma^{-n};\gamma)_{n+1}. \tag{A7}$$

To derive the last equality, we used the so-called fundamental theorem of q-calculus [29], also known as Cauchy binomial theorem, or finite q-binomial theorem [36]. Since the q-Pochhammer symbol is a product that includes factor $(1-\gamma^{-n}\gamma^n)$, we have $S_n^{(1)}=1$ for all n.

3. Calculation of sums $S_n^{(2)}$

Let us consider the following sums:

$$S_n^{(2)}=\sum_{p=0}^{n}\frac{g_{n-p}}{(\gamma;\gamma)_p}=\sum_{p=0}^{n}\frac{(-1)^{n-p}\gamma^{\frac{(n-p)(n-p-1)}{2}}}{(\gamma;\gamma)_p(\gamma;\gamma)_{(n-p)}}$$
$$=\frac{1}{(\gamma;\gamma)_n}\sum_{m=0}^{n}(-\gamma^{-1})^m\gamma^{\frac{m(m+1)}{2}}\binom{n}{m}_{\gamma}. \tag{A8}$$

Hence $S_0^{(2)}=1$, whereas, for $n\geqslant 1$, using the fundamental theorem of q-calculus,

$$S_n^{(2)}=\frac{1}{(\gamma;\gamma)_n}\prod_{m=1}^{n}(1-\gamma^{-1}\gamma^m)=0. \tag{A9}$$

Introducing the Kronecker's symbol δ_{nm}, we thus have simply

$$S_n^{(2)}=\delta_{n0}. \tag{A10}$$

4. Calculation of series $S(x)$

Here, we calculate function $S(x)$ defined by Eq. (37). Using the definition of g_n [Eq. (9)] and noting that

$$\frac{1}{x+\gamma^n}=\frac{1}{1+x}\frac{(-x;\gamma)_n}{(-x\gamma;\gamma)_n}, \tag{A11}$$

this is rewritten as

$$S(x)=\frac{1}{1+x}\sum_{n=0}^{\infty}\frac{(-\gamma)^n\gamma^{\frac{n(n-1)}{2}}}{(\gamma;\gamma)_n}\frac{(-x;\gamma)_n}{(-x\gamma;\gamma)_n}$$
$$=\frac{1}{1+x}\sum_{n=0}^{\infty}(-\gamma)^n\gamma^{\frac{n(n-1)}{2}}\frac{(-x;\gamma)_n(-x\gamma;\gamma)_n}{(\gamma;\gamma)_n(-x\gamma;\gamma)_n(-x\gamma;\gamma)_n}. \tag{A12}$$

In the second equality, we have introduced $(-x\gamma;\gamma)_n$ at numerator and denominator in order to use Jackson's transformation formula [37] (see also Identity 9 in Ref. [31]), which allows us to rewrite

$$S(x)=\frac{1}{1+x}\frac{(\gamma;\gamma)_n}{(-x\gamma;\gamma)_n}{}_2\phi_1(-x,1;-x\gamma;\gamma,\gamma), \tag{A13}$$

where ${}_2\phi_1(a,b;c;\gamma,z)$ is Heine's q-hypergeometric function [29,37], defined for $|z|<1$ by

$${}_2\phi_1(a,b;c;\gamma,z)=\sum_{n=0}^{\infty}\frac{(a;\gamma)_n(b;\gamma)_n}{(\gamma;\gamma)_n(c;\gamma)_n}z^n. \tag{A14}$$

From this definition, it appears that, because of factor $(1;\gamma)_n$, the terms of the sum ${}_2\phi_1(-x,1;-x\gamma;\gamma,\gamma)$ are all zero except for $n=0$; this term being equal to 1, we get

$$S(x)=\frac{1}{1+x}\frac{(\gamma;\gamma)_\infty}{(-x\gamma;\gamma)_\infty}=\frac{(\gamma;\gamma)_\infty}{(-x;\gamma)_\infty}. \tag{A15}$$

5. Calculation of an integral

We now calculate the integral

$$I=\int_0^\infty\frac{dP}{P}\sum_{n=0}^{\infty}g_n e^{-\beta\gamma^{-n}F_nP}, \tag{A16}$$

where the g_n are given by Eq. (9) and the F_n ($n\geqslant 0$) are positive numbers such that the sum converges.

Splitting the integral at an arbitrary $X>0$ and using Eq. (11) in the first part gives

$$I=\int_0^X\frac{dP}{P}\sum_{n=0}^{\infty}g_n\left(e^{-\beta\gamma^{-n}F_nP}-1\right)$$
$$+\int_X^\infty\frac{dP}{P}\sum_{n=0}^{\infty}g_n e^{-\beta\gamma^{-n}F_nP}. \tag{A17}$$

Then, for any $n\geqslant 0$,

$$\int_0^X\frac{dP}{P}\left(e^{-\beta\gamma^{-n}F_nP}-1\right)+\int_X^\infty\frac{dP}{P}e^{-\beta\gamma^{-n}F_nP}$$
$$=\int_0^{\beta\gamma^{-n}F_nX}\frac{dQ}{Q}(e^{-Q}-1)+\int_{\beta\gamma^{-n}F_nX}^\infty\frac{dQ}{Q}e^{-Q}$$
$$=-[E_1(\beta\gamma^{-n}F_nX)+\ln(\beta\gamma^{-n}F_nX)+\gamma_E]$$
$$+E_1(\beta\gamma^{-n}F_nX)$$
$$=-[\ln(\beta X)+\gamma_E+\ln F_n-n\ln\gamma], \tag{A18}$$

where we have introduced the E_1 function and the Euler-Mascheroni constant γ_E [38] and obtained the integrals from Ref. [38]. Using Eqs. (11) and (A4) (Appendix, Sec. 1), we finally obtain

$$I=-\sum_{n=0}^{\infty}g_n[\ln X+\gamma_E+\ln F_n]$$
$$=-\sum_{n=0}^{\infty}g_n\ln F_n-(\gamma;\gamma)_\infty\ln\gamma. \tag{A19}$$

6. Derivatives of function f_E

The first derivatives $f_E^{(n)}=d^nf_E/dx^n$ of function f_E, defined by Eq. (51), are easily calculated. Setting $y=\gamma^x$, we find that

$$f_E^{(1)}(x)=(\ln\gamma)y/(1+Ey)^2,$$
$$f_E^{(2)}(x)=(\ln\gamma)^2y(1-Ey)/(1+Ey)^3, \tag{A20}$$
$$f_E^{(3)}(x)=(\ln\gamma)^3y(1-4Ey+E^2y^2)/(1+Ey)^4.$$

This leads us to propose the general formula

$$f_E^{(n)}(x)=(\ln\gamma)^n y\frac{A_n(-Ey)}{(1+Ey)^{n+1}}, \tag{A21}$$

where A_n is the nth Eulerian polynomial [34]. This formula can be demonstrated by recurrence. Assuming that it holds for

$f_E^{(n)}$, we find that

$$f_E^{(n+1)}(x) = (\ln\gamma)^{n+1} y \frac{(1-nEy)A_n(-Ey) - Ey(1+Ey)\left.\frac{dA_n}{du}\right|_{u=-Ey}}{(1+Ey)^{n+2}}. \tag{A22}$$

The standard relation between A_n, its first derivative, and A_{n+1} [Eq. (3.3) in Ref. [34]] readily shows that Eq. (A21) also holds at order $n+1$. Since A_n is a polynomial of degree $n-1$, f_E and all its derivatives tend to zero when $x \to \infty$ and hence $y \to 0$.

7. Conditional probability for n nucleations in a time interval

Here, we demonstrate that Eq. (29) indeed gives the probability $\pi_n(T|P_0)$ for having a given number n of nucleations in a time interval of length T, conditional to the NP being P_0 at the start of the interval, and we determine the $a_p^{(n)}$ coefficients. To this end, we first derive a relation between π_n and π_{n+1}. Using the same reasoning as for π_1 and π_2 in Sec. IV A, we find

$$\pi_{n+1}(T|P_0) = \int_0^T \pi_0(\tau|P_0) P_0 e^{A b \tau} d\tau\, \pi_n(T-\tau|P_0 e^{A(b\tau-\delta)}). \tag{A23}$$

Using Eq. (23) for π_0 and assuming formula (29) for π_n, we obtain, after integration,

$$\pi_{n+1}(T|P_0) = \frac{1}{(1-\gamma)^n} \sum_{p=0}^{n} \frac{a_p^{(n)}}{1-\gamma^{p+1}} [\exp(-\gamma^{p+1}\beta E P_0) - \exp(-\beta E P_0)]. \tag{A24}$$

Identifying the factors of $\exp(-\gamma^p \beta E P_0)$ for $0 \leqslant p \leqslant n+1$ in Eq. (A24) shows that π_{n+1} indeed assumes the general form (29) provided that

$$a_p^{(n+1)} = \frac{1-\gamma}{1-\gamma^p} a_{p-1}^{(n)} \quad \text{for } 1 \leqslant p \leqslant n+1, \tag{A25}$$

$$a_0^{(n+1)} = -\sum_{q=0}^{n} \frac{1-\gamma}{1-\gamma^{q+1}} a_q^{(n)}. \tag{A26}$$

We have thereby shown by recurrence that expression (29) is valid for all n. It follows from the repeated application of Eq. (A25) for decreasing indices that

$$a_p^{(n)} = \frac{(1-\gamma)^p}{(\gamma;\gamma)_p} a_0^{(n-p)}. \tag{A27}$$

What remains to be done to determine all $\pi_n(T|P_0)$ is to find the $a_0^{(n)}$ coefficients. The first values are obtained from Eqs. (23), (26), and (28) ($n = 0,1,2$), or (28) and (A26) ($n = 4$):

$$a_0^{(0)} = 1, \quad a_0^{(1)} = -1, \quad a_0^{(2)} = \frac{\gamma}{1+\gamma}, \quad a_0^{(3)} = \frac{-\gamma^3(1-\gamma)^2}{(1-\gamma^2)(1-\gamma^3)}. \tag{A28}$$

This suggests the following general expression:

$$a_0^{(n)} = (-1)^n \gamma^{\frac{n(n-1)}{2}} \frac{(1-\gamma)^n}{(\gamma;\gamma)_n} = (1-\gamma)^n g_n. \tag{A29}$$

Assuming Eq. (A29) to hold for $0 \leqslant p \leqslant n$ and using successively Eqs. (A26) and (A27), we find

$$a_0^{(n+1)} = -\sum_{p=0}^{n} \frac{1-\gamma}{1-\gamma^{p+1}} a_p^{(n)} = -\sum_{p=0}^{n} \frac{(1-\gamma)^{p+1}}{(\gamma;\gamma)_{p+1}} a_0^{(n-p)}. \tag{A30}$$

Using Eq. (A29) to express $a_0^{(n-p)}$ for $0 \leqslant p \leqslant n$, we find

$$a_0^{(n+1)} = (1-\gamma)^{n+1} g_{n+1} S_n^{(1)}, \tag{A31}$$

where $S_n^{(1)}$ is defined by Eq. (A5) (Sec. 2). To prove that Eq. (A29) holds for any integer n, it thus suffices to verify that $S_n^{(1)} = 1$, which is done in Sec. 2. Then, using Eqs. (A27) and (A29), we obtain the general expressions of coefficients $a_p^{(n)}$:

$$a_p^{(n)} = \frac{(1-\gamma)^n}{(\gamma;\gamma)_p} g_{n-p}. \tag{A32}$$

The final expression for π_n, Eq. (30), follows directly from replacement of $a_p^{(n)}$ in Eq. (29).

8. σ^2 is bounded

Here, we show that σ^2 is bounded. Recalling that $|B_2(x)| \leqslant b_2$ for any x and that, for $0 \leqslant x < 1$, $\{x\} = x$, Eq. (57) yields

$$|R_2(E)| \leqslant b_2 J_2(E) + |K_2(E)|, \tag{A33}$$

with

$$J_2(E) = \int_0^\gamma \left[\frac{E(u+1)+2E^2u^2}{(u-1)^2(1+Eu)^2} - \frac{u+1}{(u-1)^3}\ln\frac{1+E}{1+Eu}\right] du, \tag{A34}$$

$$K_2(E) = \int_\gamma^1 B_2\left(\frac{\ln u}{\ln\gamma}\right) V_{E,2}(u)\, du. \tag{A35}$$

To obtain Eq. (A34), we noted that the first and second terms of the sum defining $V_{E,2}(u)$ in Eq. (58) are respectively positive and negative for any u, so that their difference is larger than $|V_{E,2}(u)|$. We find that

$$J_2(E) = \frac{\gamma}{1-\gamma}\left(\frac{E}{1+E\gamma} + \frac{2E}{1+E}\right) + \frac{2E}{(1+E)^2}\ln\frac{1-\gamma}{1+E\gamma} + \frac{\gamma}{(1-\gamma)^2}\ln\frac{1+E}{1+E\gamma}. \tag{A36}$$

This quantity clearly remains finite for $E \geqslant 0$ (its limit is easily calculated). In turn, we find a long analytical expression for $K_2(E)$, which also admits a finite limit:

$$\lim_{E\to\infty} K_2(E) = -\frac{\pi^2}{3(\ln\gamma)^2} + \frac{2\ln(1-\gamma)-1}{\ln\gamma} + \frac{11\gamma-1}{12(1-\gamma)} + \frac{12\,\mathrm{Li}_2(\gamma) - \gamma\ln\gamma}{6(1-\gamma)^2}, \tag{A37}$$

with Li_2 the dilogarithm function [38]. This demonstrates that, provided $\gamma < 1$, σ^2 remains finite when $E \to \infty$, in sharp contrast with the infinite increase of σ^2 in the Poisson case.

[1] K. A. Dick, Prog. Cryst. Growth Charac. Mater. **54**, 138 (2008).
[2] F. Glas, J. C. Harmand, and G. Patriarche, Phys. Rev. Lett. **104**, 135501 (2010).
[3] K.-C. Lu, K. N. Tu, W. W. Wu, L. J. Chen, B.-Y. Yoo, and N. V. Myung, Appl. Phys. Lett. **90**, 253111 (2007).
[4] Y.-C. Chou, W.-W. Wu, S.-L. Cheng, B.-Y. Yoo, N. Myung, L. J. Chen, and K. N. Tu, Nano Lett. **8**, 2194 (2008).
[5] J. A. Elliott, Y. Shibuta, H. Amara, C. Bichara, and E. C. Neyts, Nanoscale **5**, 6662 (2013).
[6] K. Reyes, P. Smereka, D. Nothern, J. M. Millunchick, S. Bietti, C. Somaschini, S. Sanguinetti, and C. Frigeri, Phys. Rev. B **87**, 165406 (2013).
[7] W. Theis and R. M. Tromp, Phys. Rev. Lett. **76**, 2770 (1996).
[8] X. Pan and X. Bao, Acc. Chem. Res. **44**, 553 (2011).
[9] A. O. Kovalchuk, A. M. Gusak, and K. N. Tu, Nano Lett. **10**, 4799 (2010).
[10] V. G. Dubrovskii, N. V. Sibirev, G. E. Cirlin, J. C. Harmand, and V. M. Ustinov, Phys. Rev. E **73**, 021603 (2006).
[11] J. Johansson, L. S. Karlsson, C. P. T. Svensson, T. Mårtensson, B. A. Wacaser, K. Deppert, L. Samuelson, and W. Seifert, Nat. Mater. **5**, 574 (2006).
[12] F. Glas, J. C. Harmand, and G. Patriarche, Phys. Rev. Lett. **99**, 146101 (2007).
[13] S. H. Oh, M. F. Chisholm, Y. Kauffmann, W. D. Kaplan, W. Luo, M. Rühle, and C. Scheu, Science **330**, 489 (2010).
[14] A. D. Gamalski, C. Ducati, and S. Hofmann, J. Phys. Chem. C **115**, 4413 (2011).
[15] C.-Y. Wen, J. Tersoff, K. Hillerich, M. C. Reuter, J. H. Park, S. Kodambaka, E. A. Stach, and F. M. Ross, Phys. Rev. Lett. **107**, 025503 (2011).
[16] V. G. Dubrovskii, N. V. Sibirev, J. C. Harmand, and F. Glas, Phys. Rev. B **78**, 235301 (2008).
[17] V. G. Dubrovskii, Phys. Rev. B **87**, 195426 (2013).
[18] V. Ruth and J. P. Hirth, J. Chem. Phys. **41**, 3139 (1964).
[19] V. G. Dubrovskii, G. E. Cirlin, I. P. Soshnikov, A. A. Tonkikh, N. V. Sibirev, Y. B. Samsonenko, and V. M. Ustinov, Phys. Rev. B **71**, 205325 (2005).
[20] P. Krogstrup, R. Popovitz-Biro, E. Johnson, M. H. Madsen, J. Nygård, and H. Shtrikman, Nano Lett. **10**, 4475 (2010).
[21] M. R. Ramdani, J. C. Harmand, F. Glas, G. Patriarche, and L. Travers, Cryst. Growth Design **13**, 91 (2013).
[22] F. Glas, M. R. Ramdani, G. Patriarche, and J.-C. Harmand, Phys. Rev. B **88**, 195304 (2013).
[23] J.-C. Harmand, F. Glas, and G. Patriarche, Phys. Rev. B **81**, 235436 (2010).
[24] F. Glas, J. Appl. Phys. **108**, 073506 (2010).
[25] I. N. Markov, *Crystal Growth for Beginners* (World Scientific, Singapore, 2003).
[26] A. D. Gamalski, J. Tersoff, R. Sharma, C. Ducati, and S. Hofmann, Nano Lett. **10**, 2972 (2010).
[27] N. V. Sibirev, M. V. Nazarenko, D. A. Zeze, and V. G. Dubrovskii, J. Cryst. Growth **401**, 51 (2014).
[28] T. Kato and J. B. McLeod, Bull. Am. Math. Soc. **77**, 891 (1971).
[29] T. Ernst, *A Comprehensive Treatment of q-calculus* (Springer, Basel, 2012).
[30] G. E. Andrews, *The Theory of Partitions* (Cambridge University Press, Cambridge, UK, 1998).
[31] J. Mc Laughlin, A. V. Sills, and P. Zimmer, Electron. J. Combin. **15**, DS15 (2008).
[32] T. M. Apostol, Am. Math. Monthly **106**, 409 (1999).
[33] V. Kac and P. Cheung, *Quantum Calculus* (Springer, New York, 2002).
[34] D. Foata, in *The Legacy of Alladi Ramakrishnan in the Mathematical Sciences*, edited by K. Alladi, J. R. Klauder, and C. Rao (Springer, New York, 2010), p. 253.
[35] N. V. Sibirev, Tech. Phys. Lett. **39**, 660 (2013).
[36] H. H. Chan, S. H. Chan, and S. Cooper, Math. Medley **33**, 2 (2006).
[37] G. Gasper and M. Rahman, *Basic Hypergeometric Series*, Encyclopedia of Mathematics and its Applications Vol. 35 (Cambridge University Press, Cambridge, UK, 1990).
[38] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).